(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
23 May 2002 (23.05.2002)

PCT

(10) International Publication Number
WO 02/41114 A2

(51) International Patent Classification[7]: G06F

(21) International Application Number: PCT/US01/46135

(22) International Filing Date: 30 October 2001 (30.10.2001)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/244,422 30 October 2000 (30.10.2000) US

(71) Applicant *(for all designated States except US)*: **RAF TECHNOLOGY, INC.** [US/US]; 16650 NE 79th Street, Suite 200, Redmond, WA 98033 (US).

(72) Inventor; and

(75) Inventor/Applicant *(for US only)*: **ROSS, David, Justin** [US/US]; 16650 NE 79th Street, Suite 200, Redmond, WA 98052 (US).

(74) Agent: **PANOFF, Christopher, V.**; Stoel Rives LLP, 900 SW Fifth, Suite 2600, Portland OR 97204 (US).

(81) Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

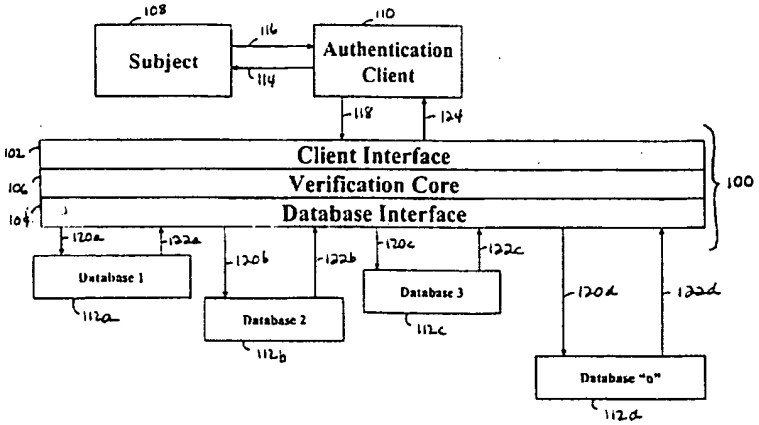
(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:
— *without international search report and to be republished upon receipt of that report*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) Title: VERIFICATION ENGINE FOR USER AUTHENTICATION

Authentication System Using a Verification Engine



(57) Abstract: An aspect of the present invention is embodied in a system for remote user authentication. An entity that wishes to authenticate a user can contact a verification engine, which, in turn, has limited access to a plurality of databases containing personal information about the user. The personal information in the databases is collected and stored by the individual operators of the databases in the ordinary course of their business with the user. The databases allow the verification engine to access the user's personal information only through predefined queries. The verification engine presents the user with the queries and the user's responses are presented to each corresponding database operator for validation. The database operators then return a confidence indication for the verification step and the verification engine combines the confidence indication from each database operator into a combined confidence indication used in authentication of the remote user.

WO 02/41114 A2

VERIFICATION ENGINE FOR USER AUTHENTICATION

Related Applications

[0001] This application claims priority from U.S. Provisional Patent App. No. 60/244,422, filed October 30, 2000, which is hereby incorporated by reference in its entirety.

Copyright Notice

Technical Field

[0003] The present invention relates to the field of remote user authentication and, in particular, user authentication employing information stored in multiple, independently controlled databases.

Background of the Invention

[0004] Presently, many systems employ security measures to protect proprietary or highly sensitive information they possess. One common aspect of the security measures often includes authentication (*i.e.*, the establishment of identity) of subjects that are using the system. As used herein, the term "subject" refers to any person or entity being authenticated. There are two primary types of authentication: user authentication and data authentication. User authentication is the process of determining whether the subject is who he claims to be. There are many possible ways of determining subject identity, and they come with varying degrees of security. Traditionally, subject authentication has been done in three ways: recognition of the subject, shared knowledge, and possession of a token by the subject.

[0005] When implementing shared knowledge for subject authentication, the system knows things about the subject that should not be common knowledge. When the subject demonstrates that he also knows these things, identification is achieved. The information typically used in the authentication procedures can be divided into two types: in-wallet data and out-of-wallet data. In-wallet data is information you know about yourself or can readily put your hand to. It includes your name, address, telephone number, drivers license number, social security number, checking account number, credit card numbers, mother's maiden name, and the like. This data forms the basis for some of the simpler authentication systems of the prior art.

[0006] Out-of-wallet data is information about you that would take you a little effort to find out, but that you probably have in your filing system or somewhere equally accessible with some effort. It includes information such as the amount of the last transaction with your checkbook or credit card, the holder and amount of your mortgage, your credit rating, your bank balance, and the like. Incorporating out-of-wallet information into an authentication system is more complicated, and thus, it is found less frequently in typical systems.

[0007] Passwords are a form of both shared knowledge and of tokens. For a password to work properly, both sides must know it, and only both sides must know it. Shared knowledge includes some nonobvious information as well. The billing address on your credit card, for example, is known to you, of course, but is accessible from the card issuer by any merchant who takes that credit card.

[0008] When performing a transaction, it is often necessary to fill out a form containing information important for the transaction. Quite apart from the need to authenticate the subject, the data in those forms must also be authenticated. Data authentication consists of applying business rules to the forms at the time of capture. This is done at the point of recognition for paper forms, and while the subject is online, in the case of Web forms. Data authentication can be as simple as ensuring a column of numbers adds up, or as complicated as verifying several associated tax forms to be sure a particular number was both correctly calculated and correctly copied.

[0009] In addition to attempting to overcome technological hurdles, typical authentication systems of the prior art must frequently overcome social hurdles as well. Many people are concerned with the growing availability of personal information on the

Internet. There are equal concerns around too much information being stored in one place. This is especially true when the information is being stored by a governmental entity, raising "big brother" concerns. An authentication system based on shared information is more convenient and less expensive for the subject, but it is more invasive of privacy. A system based on shared information can only work when based on information that only the subject and the authenticator knows. Because the subject and the authenticator have no contact other than the current transaction, this means that the shared information must be private information about the subject that the authenticator also possesses.

[0010] Although everyone would prefer to keep private information in the hands of the subject, the alternative to using it is identity theft. A system that works entirely online, if denied access to private information, will be exposed to identity theft. One simply cannot prevent identity theft without some form of authentication. It is also clear that the security of an authentication system and its obtrusiveness are inversely related - the more secure the system, the harder it is to make it unobtrusive. This tension inherent in authentication processes illustrates one of the limitations of proposals such as the "national identity card." There are various proposed uses for such a card, but primary among them are prevention of identity theft, limiting social benefits to legal residents, tracking those derelict in child support and other payments, fighting drug trafficking, providing identification for travelers, and making it easier to track those in this country legally or illegally.

[0011] Each of these is a legitimate social concern. Nonetheless, any system based on a single card allowing access to a centralized source of identity-establishing information suffers from two important drawbacks. The first, is establishing that a presented card is itself legitimate. Although states recently have taken measures to protect their driver's licenses and ID cards from forgery, fake ID cards are a long-standing problem. It is, for example, very difficult to establish from an ID card whether two people who look alike are actually distinct individuals.

[0012] The security needs of both the nation and the individual citizens must always be balanced against the requirement to preserve a free and open society. Authentication system, as viewed on the whole, should not be deemed overly intrusive.

Summary of the Invention

[0013] The present invention provides a method and system for authentication that are more secure, less intrusive, more flexible, and easier to keep current than the prior art.

This is accomplished through strategically employing a verification engine for authentication procedures.

[0014] The engine accepts personal data from a subject being authenticated. Typically, the data can be collected from the subject as part of a financial transaction with an agency of the federal government, financial institution, or commerce entity, or incident to some security procedure, etc. As used throughout these systems and the attached claims, the entity requesting authentication of the subject is termed the "authentication client" or "client." The data collected from the subject can then be compared to information contained in independent databases. Each database queried returns a confidence rating indicating how well the data matches. The verification engine combines these rating and returns them to the client, which can apply business rules of its own to decide whether to accept the subject as a valid user. Fuzzy logic algorithms can be used to determine the various confidence levels.

[0015] There are three main components to the verification engine. The first component controls access to the overall system by users. It provides a single interface, with a single set of commands, to all system clients. While the interface is standardized, it can still be customized based on the particular needs of each client. The second component interfaces with each remote database, speaking the language required by that database, both for querying and for receiving back the answers to queries. The third component, the verification core, operates in connection with the other two main components. It recognizes the queries sent to it by clients, because the queries are predefined based on the needs of the individual client. The verification core also possesses the protocols to translate those queries into queries to the individual databases, and it can assemble their responses into a coordinated response to the client.

[0016] The verification engine tightly controls the questions it asks of each independent database. This enables it to meet the stringent privacy requirements of many database holders, thus encouraging the use of their information in legitimate ways not previously allowed by the database operators. The verification engine also ties into a broad range of information providers. Because the verification engine has access to multiple databases that were previously unconnected, it can answer questions that cannot be answered by any single database. Because the verification engine controls access to those databases in ways established through agreement with the database owners, it overcomes the reluctance of

many database owners to allow access to their information. Thus, it can include various databases previously unavailable for general use.

[0017] Information is stored where it has a legitimate reason to be stored: with the database owners that acquired the information as part of their ordinary businesses or affairs with the subject. The verification engine allows legitimate access to personal data concerning a subject being authenticated, but it keeps others from browsing. Authentication clients are only licensed for specific predefined queries. Queries designed to browse database records or "read-out" information are not enabled by the verification engine.

[0018] Additional aspects and advantages of this invention will be apparent from the following detailed description of preferred embodiments thereof, which proceeds with reference to the accompanying drawings.

Brief Description of the Drawings

[0019] Figure 1 is a schematic representation of an authentication system employing a verification engine consistent with the present invention.

[0020] Figure 2 depicts a specific embodiment of the present invention in the context of online authentication for an e-commerce transaction.

Detailed Description of a Preferred Embodiment

[0021] Embodiments consistent with the present invention provide the ability to authenticate a subject even if the subject has had no prior contact with the system and has had no access to digital signatures, certificates, or passwords. One example of such a subject would be a person who is ready to retire and wants to set up his Social Security payments. Such a subject would possess or have the ability to access a computer and a network connection, such as through the Internet, but he is presumed to be an otherwise unsophisticated user. Another example of a user requiring authentication would be a subject trying to effect an online payment at an online store, or a subject trying to effect a transaction with his financial institution. The problem is to authenticate these subjects without requiring undue effort on the part of the subjects. To facilitate clarity and for ease of discussion, the references to the subject herein will assume a remote individual or entity that is accessing the authentication system through a network, such as the Internet. It is equally within the scope of the present invention to conduct live, real-time authentication of the subject in-person.

[0022] A preferred embodiment of the present invention implements a verification engine for authenticating a subject using predefined queries requesting personal identifying information from the subject. The identifying information comes from multiple third-party databases that have gathered that information in the ordinary course of their business or other relationships and dealings with the subject. The engine itself can adopt a client-server architecture and can be modeled after systems such as those used with market success in the postal and form recognition fields. The flexible design can enable processing of queries at rates up to 850 per minute and is scalable as demand on the system increases.

[0023] There are three main parts to the verification engine. The first part controls access to the overall system by authentication clients. It provides a single user interface, with a single set of commands, to all system users. The queries it can present, as discussed in detail below, however, can be tailored to the particular needs of a specific client. The second part interfaces with each remote database, speaking the language required by that database, both for querying and for receiving back the answers to queries.

[0024] The third part, functions in the middle, connecting the other two main parts. This part is termed the "verification core." It understands the queries sent to it by users. It knows how to translate those queries into queries for the individual databases. The verification core knows what questions each connected database can answer, how likely it is to provide an answer, how long it generally takes to do so, and it possesses functioning protocols for querying each database. The verification core can also assemble the responses from the various independent databases into a coordinated response to the authentication client. The verification core obtains this information through collaboration with the independent database operators to develop the specific queries that will be allowed for use in the authentication procedure.

[0025] Figure 1 illustrates an authentication system incorporating a verification engine consistent with the present invention. With respect to Figure 1, the verification engine 100 includes various components which include a client interface 102, a database interface 104, and the verification core 106. A subject 108 that either gives or requires authentication communicates with an authentication client 110. The subject 108 can provide identifying information to the authentication client 110, and the subject 108 can also answer queries posed by the authentication client 110. The authentication client 110 gives and receives information to and from the verification core 106 through the client interface 102. The

verification core can then provide the information identifying the subject 108 to one or more independent, third-party databases 112a through 112d for authentication. Database "n" 112d illustrates that a potentially unlimited number of independent databases can be provided for authentication with the verification engine.

[0026] Continuing with Figure 1, after the subject 108 has identified himself to the authentication client 110, the authentication client 110 can present a predefined query 114 to the subject. This predefined query has been licensed for a specific use by this authentication client 110. The actual question and scope of the query depend on the authentication services being provided by the authentication client 110. The subject 108 then returns a response to the query 116. The authentication client 110 then forwards 118 the response to the verification engine 100 for authenticating the subject 108. The verification engine 100 then transmits 120a through 120d the response from the subject 108 to multiple of the databases 112a through 112d. Each database 112a through 112d that receives the information checks it against the identifying information it stores for the subject 108 and returns a confidence indication 122a through 122d to the verification engine. The verification engine 100 combines the individual confidence indications 122a through 122d into a combined confidence indication 124 that is provided to the authentication client 110 for authenticating the subject 108.

[0027] On the front end, the verification engine can establishes the identity of its authentication client with standard digital certificates, passwords, and user names. Each authentication client has a specified list of questions it can ask the system. For example, an airport security guard (in possession of a passenger's driver's license) may only ask such a system "Does John Smith live at 123 Main Street, Jackson, Mississippi, and have Mississippi Driver's License number 549-34-2218?" The only response he can receive back is either a yes or a no, the confidence the system has in that answer, and possibly a picture of the legitimate holder of the license. He cannot, for example, ask: "who lives at 123 Main Street?"

[0028] If an FBI investigator is the authentication client, on the other hand, he may be authorized for considerably greater access to the system. He may, for example, be able to ask "Give me the passport numbers of all males of Bahrain residency who entered the United States through New York or Boston between May 6 and June 7, 2001 on British Air or United Airlines flights from Paris or Zurich." The scope of permissible questioning

would depend on the permissions the operator of the databases being asked to authenticate the responses would be willing to authorize. Because the system controls access by limiting the kinds of questions each type of user can ask, and because there is no central repository of the data, the system is much more protective of privacy.

[0029] On the back end, the system also has tight controls for the questions it asks of each database. This enables it to meet the stringent privacy requirements of many database holders. The ability to control the types of queries being made against the database provides many database providers the incentive, or at least comfort level to make their database available to the verification engine. The independent database operator can authorize a query to access whatever level of information they are comfortable with. This makes available the use of their information in legitimate ways that were not previously possible.

[0030] The verification engine system is designed to be both secure and unobtrusive. In a preferred embodiment, the only data the verification engine will have direct access to is certain low-sensitivity in-wallet data. This information can come from several commercial sources. The data includes name, address, telephone number, and other personal information that most people do not consider particularly sensitive, but that help the engine authenticate a subject at the simplest level. The information can be obtained from the subject as part of a Web form when a person first uses the system. Asking for it as part of a Web form takes advantage of the habit of filling such information into forms. If done skillfully, most subjects will not even know we are using that information for authentication.

[0031] The more sensitive out-of-wallet data, such as creditworthiness and credit card information, remains in the hands of companies that naturally hold that information. Although the verification engine will know the results of the queries, the information itself is never directly accessible by the verification engine or the authentication client. The verification engine simply provides a gateway to the information, thus offering a workable compromise between authentication and privacy. The system also meets the requirements of many governmental agencies by allowing access to personal data under tightly controlled conditions for legal, social, or medical needs.

[0032] The user information and the specific databases being used for authentication can vary widely. This affords the present system a significant amount of flexibility. One

perfectly suitable commercial database available to the verification engine is the Crystal Database offered by RAF Technology, Inc. of Redmond Washington. Compiled from a wide range of reliable sources including the US Postal Service, this database is condensed into data crystals that take a fraction of the memory of the original and can be rapidly accessed. Crystallizing the data also obscures the information so that it cannot be read. As a result, the Crystal Database data crystals can be installed on the verification engine hardware without actually making the data visible to anyone. Access to the Crystal Database is strictly via predefined queries. In effect, the Crystal Database is lent, not given.

[0033] Other databases of personally identifying information can be obtained from such sources as VISANET, the Social Security Administration, the Internal Revenue Service, various of the national credit bureaus, national telephone directory services, etc. State-level sources of information can also be queried depending on the transaction involved. Examples include the Department of Motor Vehicles, the various taxation offices, etc.

[0034] Because there are multiple ways to establish identity, Security or law enforcement personnel can use the personal information databases in different, and unpredictable ways. Because the information is stored in multiple unrelated databases, it becomes extremely difficult for an identity thief or other criminal to place false data in all of them. Among the information and combinations it uses to establish identity are: name and aliases plus address; name plus telephone number(s); Social Security Number or other federal identifier; passport number; driver's license number; name plus mother's maiden name; name plus patronymic; bank account or other financial information; green card number; resident alien number, as well as many others.

[0035] Figure 2 illustrates the operation of the verification engine in an authentication system provided for authenticating a subject in an electronic commerce transaction. In this context, the subject is labeled the "Customer," the authentication client is the "e-commerce site," the independent databases are the "trusted Validator," and the verification engine is being operated by the "Authentex" entity. For simplicity in Figure 2, the queries and response paths are illustrated as going directly to the verification engine, rather than through the authentication client.

[0036] With reference to Figure 2, the Customer (box 1) logs onto an e-commerce site (box 15) for which Authentex (box 10) provides authentication. The system will ask the

Customer a series of appropriate questions (box 8) to authenticate his identity. These questions center on in-wallet data that Authentex itself possesses, and out-of-wallet data possessed by a trusted Validator (box 3) such as a bank or credit bureau. Authentex holds in-wallet data and provides the gateway to Validators who hold out-of-wallet data.

[0037] The questions are "appropriate" in that they fit the situation. Clearly asking for name, address, phone number, and Social Security Number would be seeking appropriate in-wallet data that can be used to authenticate the Customer. Choosing appropriate out-of-wallet questions is more difficult. Out-of-wallet data (box 4) and physical validation of the Customer (box 2) were collected by the Validator through the course of its normal interactions with the Customer, independent of any connection with Authentex. The Validators build up a database of information, and a series of queries that can be put to that database. Authentex and the Validator establish a set of allowed queries (box 5) which is a subset of all the queries permitted by the Validator's database, chosen to provide proper authentication while being as unobtrusive as possible. Effectively, the Validator is digitally vouching for the Customer.

[0038] Choosing what queries are appropriate can be quite a challenge, and the queries can change with the nature of the requested transaction. If the Customer wants to set up automatic payments to his savings account, an appropriate question would be "what is your savings account number." Asking for his credit card number would be inappropriate.

[0039] Both in-wallet and out-of-wallet questions (box 8) are presented as items to be filled out in a Web form. This is far less obtrusive than direct questions, and seems much more natural to the Customer. In fact, if done skillfully, the Customer will never know he is being authenticated at all. The answers (box 9) to in-wallet queries are checked (box 12) against Authentex's Personal Information Database (box 11) and a confidence in the answers (box 13) returned to Authentex. The answers (box 9) to out-of-wallet queries are checked (box 6) against the Validator's database and a confidence in the answers (box 7) is returned to Authentex. Authentex assembles the answers and, using fuzzy logic, determines and passes on its overall confidence (box 14) to the e-commerce site, which makes the final authentication decision.

[0040] It should be noted that the verification engine can simultaneously serve multiple authentication clients. Similar, each authentication client employ the verification engine to authenticate multiple subjects.

[0041] The present invention also provides a suitable solution to the "national identity card" debate. The verification engine fulfills the legitimate needs of a national identity system and information bank, while meeting the privacy concerns of individual citizens. It meets the needs of citizens by leaving their personal data in the possession of those whom they perceive as having legitimate need for open access to it.

[0042] It will be obvious to those having skill in the art that many changes may be made to the details of the above-described embodiments of this invention without departing from the underlying principles thereof. The scope of the present invention should, therefore, be determined only by the following claims.

Claims

1. A user authentication system comprising:

an authentication client for requesting authentication of a subject;

a user interface to receive the authentication request from the authentication client;

multiple independently operated databases, each database storing information associated with the subject, the associated information being accessible through predefined queries to identify the subject; and

a verification engine for facilitating authentication of the subject by receiving the authentication request, selecting one or more of the predefined queries, presenting the one or more selected queries to the subject via the authenticating client, receiving from the subject an answer to each of the one or more selected queries, and presenting the answer to the multiple independently operated databases for a validation response.

2. The system of claim 1 wherein the associated information in the multiple independently operated databases includes out-of-wallet data identifying the subject.

3. The system of claim 1 further comprising a personal information database coupled to the verification engine, the personal information database containing in-wallet data identifying the subject.

4. An authentication system comprising:

an authentication client for desiring authentication of an authentication subject;

a plurality of independent database systems storing information identifying the authentication subject, the identifying information being accessible through predefined queries; and

a verification engine to receive from the authentication subject, via the authentication client, an answer to each of the predefined queries, to obtain from each of the plurality of independent database systems a corresponding authentication confidence for each answer, and to combine the corresponding authentication confidence for each answer into a combined authentication confidence.

5. A user authorization method comprising the steps of:

presenting to an authentication subject one or more predefined queries from each of multiple independent databases of identifying information;

receiving from the authentication subject an answer to each of the selected queries;

presenting each answer to at least one of the multiple independent databases that has corresponding identifying information;

obtaining from the multiple independent databases an authentication confidence level for each answer; and

combining the authentication confidence level for each answer into a combined confidence level for authenticating the authentication subject.

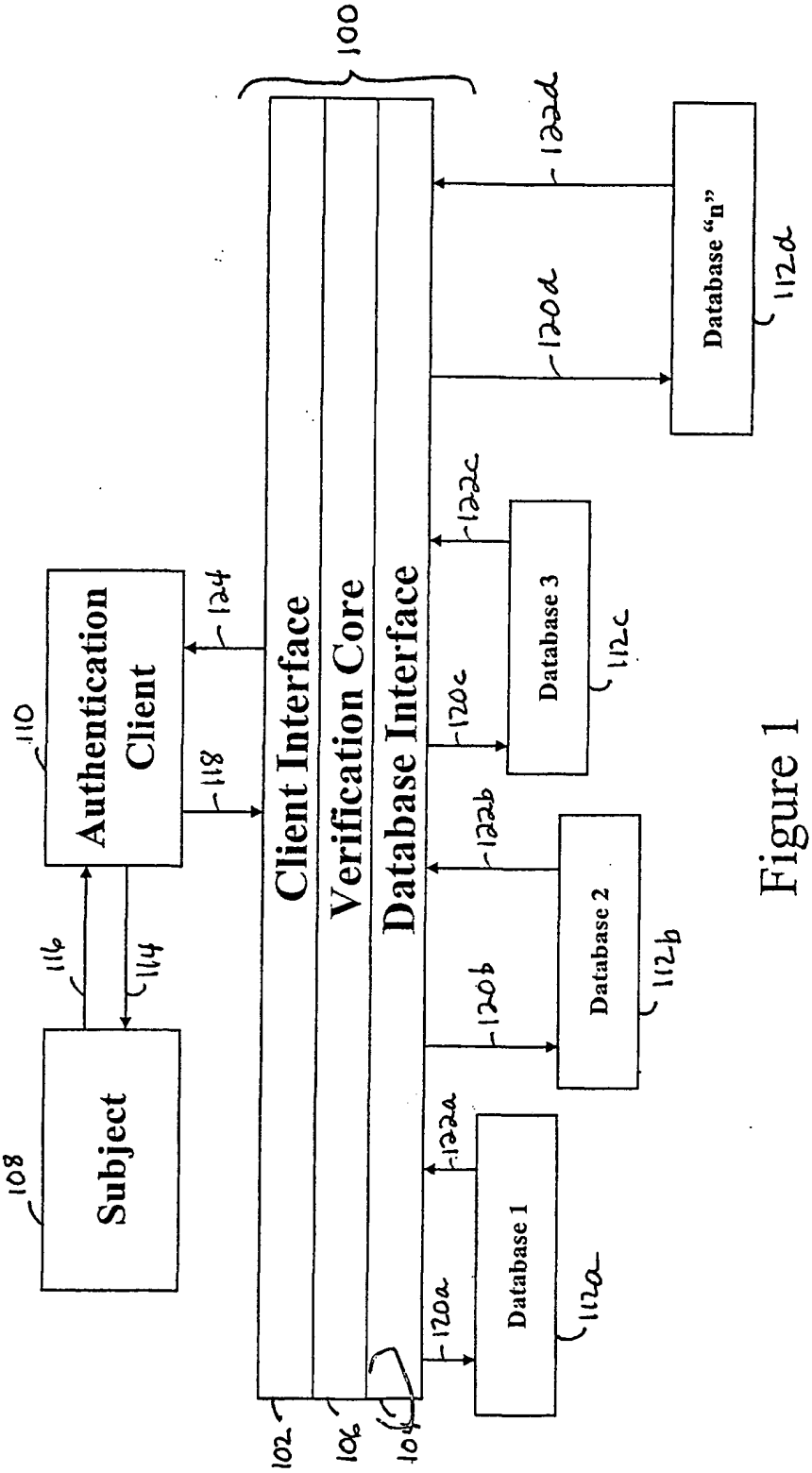

Authentication System
Using a Verification Engine
100
Client Interface
Verification Core
Database Interface
102
104
106
Subject
108
Authentication Client
110
114
116
118
124
120a
120b
120c
120d
122a
122b
122c
122d
Database 1
112a
Database 2
112b
Database 3
112c
Database "n"
112d

Figure 1

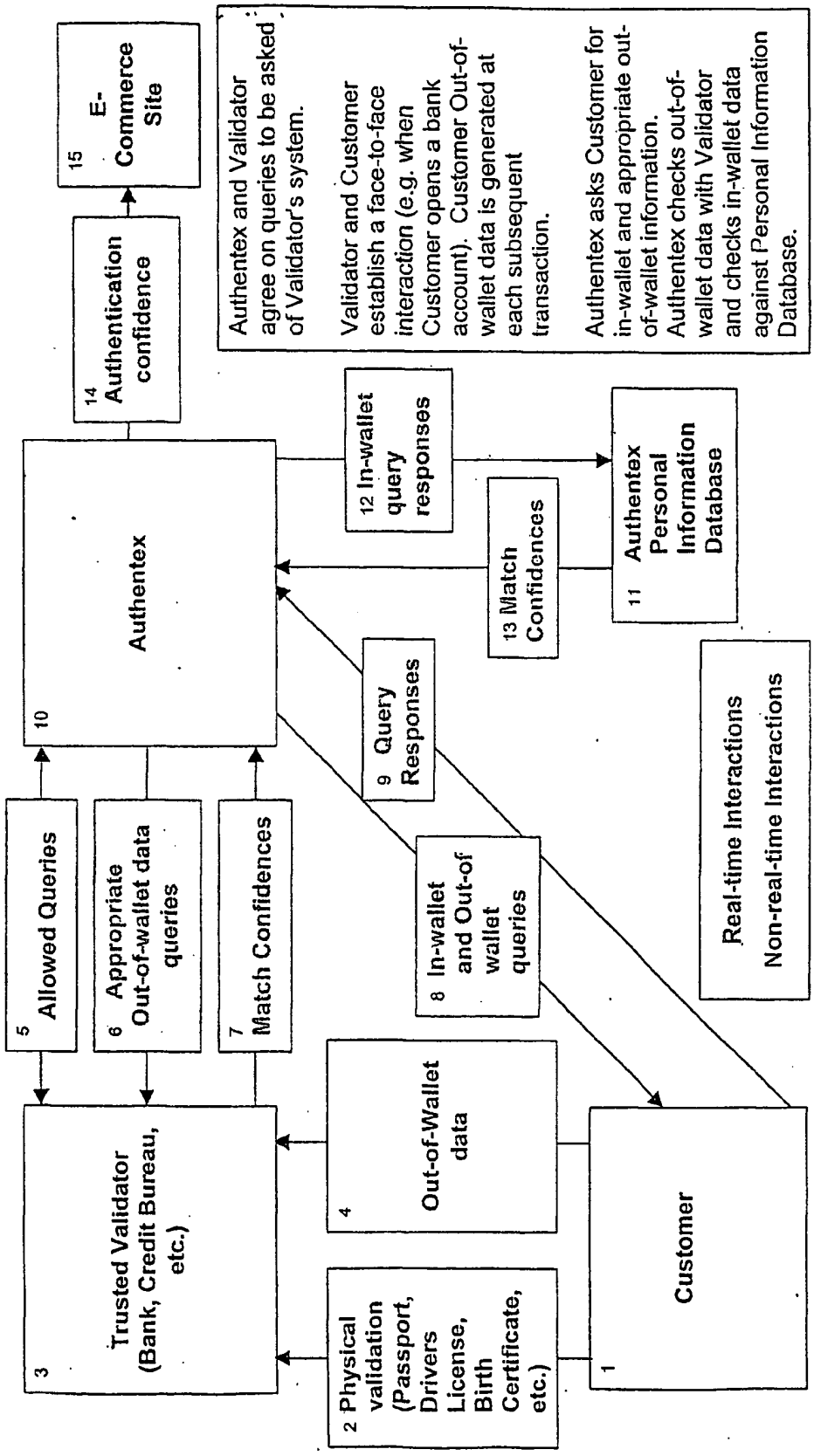

1 Customer
2 Physical validation (Passport, Drivers License, Birth Certificate, etc.)
3 Trusted Validator (Bank, Credit Bureau, etc.)
4 Out-of-Wallet data
5 Allowed Queries
6 Appropriate Out-of-wallet data queries
7 Match Confidences
8 In-wallet and Out-of wallet queries
9 Query Responses
10 Authentex
11 Authentex Personal Information Database
12 In-wallet query responses
13 Match Confidences
14 Authentication confidence
15 E-Commerce Site
Authentex and Validator agree on queries to be asked of Validator's system.
Validator and Customer establish a face-to-face interaction (e.g. when Customer opens a bank account). Customer Out-of-wallet data is generated at each subsequent transaction.
Authentex asks Customer for in-wallet and appropriate out-of-wallet information. Authentex checks out-of-wallet data with Validator and checks in-wallet data against Personal Information Database.
Real-time Interactions
Non-real-time Interactions

Figure 2